DOCUMENT RESUME

ED 079 240 SP 006 632

AUTHOR	Rothney, John W. M.
TITLE	Evaluating and Reporting Pupil Progress. What Research Says to the Teacher Series. No. 7.
INSTITUTION	National Education Association, Washington, D.C. Association of Classroom Teachers.
PUB DATE	72
NOTE	36p.
AVAILABLE FROM	National Education Association, 1201 Sixteenth St., N.W., Washington, D.C. 20036 (Stock No. 387-11812, No price quoted)
EDRS PRICE	MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	*Academic Achievement; Individual Characteristics; Personality Tests; *Psychological Evaluation; *Student Evaluation; Test Construction; *Tests; *Test Validity

ABSTRACT

This booklet, one in a series entitled "What Research Says to the Teacher", deals with the evaluation and reporting of pupil progress. An introduction discusses briefly the purposes, criteria, and methods of student evaluation. Section I examines the standard criteria for student achievement: school marks, standardized tests, curves, essay tests, and teacher-made objective tests. A second section asks questions concerning the personal-social development of the student. The concept of using checklists and rating scales to measure such variables as character and citizenship is discussed, as are other forms of personality rating and character description. Section III considers the proper time for evaluation, as a categorizing method or as a diagnosis. Also briefly discussed are report cards, cumulative records, and teacher attitudes toward reporting student progress. (JB)

FILMED FROM BEST AVAILABLE COPY

WHAT RESEARCH SAYS TO THE TEACHER

7

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

NEA

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

U.S. DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Evaluating and Reporting Pupil Progress

John W. M. Rothney

ED 079240

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to:

In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

ASSOCIATION OF CLASSROOM TEACHERS
NATIONAL EDUCATION ASSOCIATION

Copyright © 1955, 1972
National Education Association of the United States
NEA Stock No. 387-11812

Evaluating and Reporting Pupil Progress

CONTENTS

EXPLANATION 2

STUDY YOUR OWN PROGRAM 3

PURPOSES OF EVALUATION 4

EVALUATION IN TERMS OF OBJECTIVES OF EDUCATION 5

WHAT SHOULD WE EVALUATE? 5

HOW SHOULD WE EVALUATE? 7

ASSESSMENT OF SCHOOL ACHIEVEMENT 7
- Limitations of School Marks 8
- Improving Marking 9
- Standardized Tests 10
- Use of Norms and Standards 11
- Purpose of Essay Tests 12
- Teacher-Made Objective Tests 13

ASSESSING PERSONAL-SOCIAL DEVELOPMENT 15
- Many Questions Unanswered 15
- Checklists and Rating Scales 15
- Observations vs. Measurement 17
- Use of Anecdotal Records 17
- Limitations of Anecdotal Records 19
- Description vs. Rating 21
- Personality and Interest Questionnaires 22
- Personal Documents and Projective Technics 23
- Not in Isolation 25

WHEN SHOULD WE EVALUATE? 25
- Evaluation, A Year-Round Procedure 26
- Evaluation as Diagnosis 26

Table of Contents continued on page 2

WHO SHOULD EVALUATE? 26

REPORTING PUPIL PROGRESS 27
Typical Report Cards 27
Cumulative Records 28
Teachers' Attitudes Toward Reporting Pupil Progress 29
Teacher-Parent Conferences 29
New Reporting Procedures Needed 30

A FINAL WORD 31

SELECTED RESEARCH REFERENCES 31

EXPLANATION

The author has tried to select from the research material on evaluating and reporting pupil progress the items which promise most to be of help to the classroom teacher. It is not a complete summary of research, but a statement of some of the practical implications of research. The recommendations made in this booklet are those which the author, John W. M. Rothney of the University of Wisconsin, believes to be soundly supported by research. His original manuscript was reviewed by W. Earl Armstrong, National Council for Accreditation of Teacher Education; William G. Brink, Northwestern University; and John W. Shreve, Cincinnati Public Schools. Changes were made by the author on the basis of the suggestions of the reviewers and of the staff of the NEA Research Division.

EVALUATING AND REPORTING PUPIL PROGRESS

To MANY teachers the process of evaluating and reporting pupil progress is one of giving a series of exercises and tests, marking them, adding or averaging the marks, and entering them on a small card which is to be taken home to be signed by a parent, usually the mother. The process is often quick, simple, and terminal. Research has shown us, however, that if evaluating and reporting pupil progress is to be effective, it must be a continuous, cooperative, and cumulative procedure.

STUDY YOUR OWN PROGRAM

While you are reading the following pages, you might consider your own situation, look at your own procedures, and consider what you might answer to the following questions:

1. Are you using obsolete (once useful) methods of evaluating and reporting? Most of these would die gracefully but for the inertia of a few school people and parents who want them.
2. Have you clarified the terms you use in evaluating and reporting? Consider in answering this one whether your marks and ratings are really clear and meaningful both to children and to adults.
3. Are some of your procedures among those so badly designed that they should never have been born? Are you still using some that are now recognized as incorrect? Do they continue to live because of the psychological innocence of some users? You may be trying to evaluate too many characteristics too often and too superficially. From this angle consider the personality "tests," and even some "objective" tests of intelligence and achievement.
4. Are you using procedures with *every* pupil that should be used for a few special cases? Consider in answering this one whether or not your evaluating and reporting procedures put great emphasis on the traits found only in a few exceptional

students. Are you using devices that force you to put 9 out of 10 pupils in an "average" category? Has the "problem child" stolen the show?

5. Do you use modern well-proven devices and are you alert to revisions and improvements in these as they are developed? Consider in answering this one whether or not you keep up with recent developments in cumulative records, anecdotal records, behavior descriptions, and reporting procedures to the parents.

6. Do you have a continuing program of experimentation with frankly experimental evaluating and reporting techniques and are you cautious in using them until their value can be demonstrated? Consider here such things as the new "culture-free" tests, projective devices, sociograms, and pupil participation in evaluation and reporting.

PURPOSES OF EVALUATION

Evaluating and reporting processes are essential steps in the current guidance, future guidance, and transfer of pupils but they serve other purposes as well. They may confirm estimates of the effectiveness of the teaching that is done and cause us to doubt the value of our instructional methods and materials enough so that we may undertake their revision or reject them as ineffective. The results obtained by evaluation may provide classroom teachers and parents with such psychological security that they become strong in their support of the schools. They may also spur us on to make improvements that are found to be desirable. Results of evaluative procedures may enable teachers to respond effectively to critics of the schools and to enlist their cooperation in further developments. In short, if classroom teachers have a sound system of evaluating and reporting pupil progress, both the pupils and the public will know more about school objectives and the progress being made toward those objectives. Such information also helps the classroom teacher himself.

The results of evaluating pupil progress can be reported in many different ways. Pupil development can be indicated by scores or letter grades, by oral or written statements, or by any combination of such methods. Although it does not seem desirable

to let the form of the report determine the kind of evaluation devices that are to be used, or vice versa, the processes are closely related. They are separated as far as possible in the discussion that follows simply for convenience and emphasis. We shall begin with a review of evaluative techniques and turn later to ways of reporting progress.

The chief problems in this area revolve around the four questions: What should we evaluate? How should we evaluate? When should we evaluate? and Who should evaluate?

EVALUATION IN TERMS OF OBJECTIVES OF EDUCATION

It is customary to talk about evaluating pupil progress in terms of the carefully selected objectives to be achieved. We say that objectives should be stated so that they can be evaluated; the evaluations should be of such a character as to include the objectives as stated.

Suppose, for example, we want to determine whether pupils have improved in citizenship. It will be necessary to define "citizenship" by breaking it down into particular items about which specific evidence of pupil progress can be obtained. Both the evaluation of, and the report on, progress in citizenship will not be in terms of a mark that was earned but in the language that was used in defining the objectives. We will not say then that John gets an "A" in citizenship, but we will say that he has carried out specific activities such as doing his share in keeping the room clean, serving on the traffic patrol, or participating in groups concerned with social action.

If one thinks about the problem of getting adequate information about pupil progress in just this area of citizenship, it will be seen that the "what, how, when, and who" questions given above are all raised separately and in combination and that answers to them are difficult to obtain.

WHAT SHOULD WE EVALUATE?

If the purpose of our instruction is to produce certain changes in pupils, we also should find out whether those changes are

taking place. "But," you say, "there are so many objectives that evaluation of them would be impossible." What then should be done? Should we reduce the number of objectives, should we lump the common ones for evaluation purposes, or should we just attempt to evaluate some of them and have faith that the others are being accomplished?

Usually, when school staffs analyze their educational objectives, they find that many of the items listed have much in common. Long lists of objectives often can be reduced to such major classifications as these:

1. The development of effective methods of thinking
2. The cultivation of useful work habits and study skills
3. The inculcation of constructive social attitudes
4. The acquisition of a wide range of significant interests
5. The development of increased appreciation of music, art, literature, and other esthetic experiences
6. The development of social sensitivity
7. The development of better personal-social adjustment
8. The development of skill in effective communication
9. The acquisition of important information
10. The development of physical health
11. The development of a consistent philosophy of life.

After these major objectives have been set up, the next steps consist of trying to get specific evidence of pupils' behavior with respect to each of these categories. We seek to identify situations in which such behavior can be observed or measured. Finally, we try promising methods for interpreting evidence and making meaningful statements about the progress of pupils.

In brief, you will decide what important changes you are trying to produce in pupils. You will state how those changes may be expected to be shown in pupils' behavior and you will measure the changes that take place. You then study and interpret the observed changes so that your reports will be meaningful to the pupil himself, to his parents, and to anyone who wants to know what progress a particular pupil is making.

If classroom teachers are convinced that the goals they set are worthwhile, they must try to evaluate progress toward them. Failure to attempt systematic evaluation results in our listing too many undefined goals and attempting to justify teaching

on the bases of boldness and faith. *Neither* of these situations is good for pupils, parents, or the public; neither is good for education.

HOW SHOULD WE EVALUATE?

When many people think of evaluation, they think of tests. Sixty million standardized tests were given to 20 million persons during one recent year. Many of those tests were not labeled "achievement" tests, but most of them did measure achievement from which aptitude was inferred.

The main purpose of testing is not to grade or rank pupils but to assist classroom teachers in getting evidence of achievement of growth. Specialists in measurement have largely failed in constructing standardized tests to measure the totality of behavior: The traditional mathematics and logic which they use are handmaidens in the science of a piecemeal mechanical view of the world. They have not yet produced total understanding of the human personality which the classroom teacher must have.

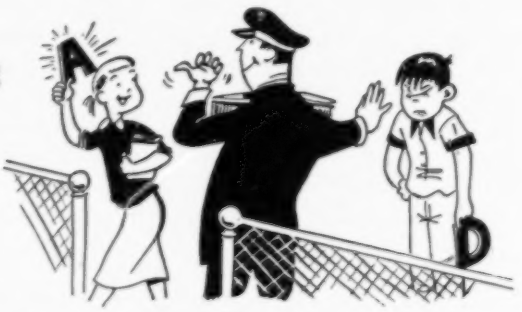
The period when such procedures were uppermost and in which quotients of intelligence and achievement were computed for most children seems now to be drawing to a close. It has served its purpose in making classroom teachers more aware than previously of individual differences among pupils. It has also made them aware of the limitations of the evaluative devices that have been used. We are now ready to pass from "the quotient stage" to one in which standardized tests will provide a small though still important place in evaluation programs. The days of mass testing seem to be numbered along with those of the "complete battery of tests," imposed on pupils and teachers. Instead we are using selective and differential testing. Classroom teachers are learning how and when to fit well-chosen tests into a carefully planned and flexible program of determining the progress of particular pupils.

ASSESSMENT OF SCHOOL ACHIEVEMENT

There is general agreement that, although teachers' marks are often unreliable and invalid indexes of growth, they are indis-

pensable tools. Marks are the coin of the school realm. They continue to be the measures of school success—the keys that open doors of educational institutions for entrance and for exit. The marks given by some classroom teachers have great value. Others may mean nothing more than that the student has official permission to forget what he has learned. In either case they seem likely to continue to be the principal basis for honor awards, promotion, and placement in schools. For a long time to come parents will accept them as the basic evaluative device.

High marks open many doors—



The many suggested substitutes for marks, such as home visits, descriptive reports, and interpretations of test scores, have not yet been proven and perhaps cannot be proven by rigorous experiments to be better than marks. All of them require more time than school people are always willing and able to give to evaluating and reporting pupil progress. Since, then, marks are likely to be with us for some time, classroom teachers will want to recognize the limitations of school marks and to examine methods that may be used to improve them. While doing so, they will be experimenting with procedures which may eventually make obsolete the school mark as we now know it.

Limitations of School Marks

A single mark cannot indicate to a pupil the points on which he needs to improve. Marks indicate no next steps for him, his parents, or his future teacher. They are simply the judgment of

the teacher, possibly affected by unrelated matters. They usually average out judgments about various elements in a pupil's progress so that meaning and value are lost. They often cause harm by increasing senseless competitiveness among pupils.

School marks may direct the attention of pupils, parents, and classroom teachers away from the *real* purposes of education toward symbols that represent success but do not emphasize its elements or meaning. They frequently permit and encourage the calculation of a meaningless rank in class or a composite score. The academic honor roll, which has been questioned by some investigators, tends to be continued by the school mark system. Studies show that higher marks are likely to be given by women than by men teachers and that both tend to give better grades to girls and to pupils who come from high socio-economic levels. These practices cause inequities.

Improving Marking

Recognition of the limitations of marks may suggest some of the things that classroom teachers can do even when required to prepare school marks for the office file. First, define very clearly what each mark means in terms of pupil *development* using descriptions of activities other than, or at least in addition to, mere repetition of memorized material. Classroom teachers may enlist the aid of pupils and parents in defining the kind of development that is expected. Then pupil growth in those activities can be observed from landmarks estimated at the beginning of the period of instruction. We should always be mindful, while doing so, that growth rates vary from pupil to pupil.

As the school year proceeds, the classroom teacher helps each pupil to recognize the points on which he needs to improve, and suggests next steps to him and to his parents. Teachers will examine their own biases and prejudices for or against particular kinds of pupils and guard against them in making their appraisals. They will not average the entries in their records so that evidence of particular strengths and weaknesses will be lost. They will make it clear by word and action that their classrooms can be *cooperative* as well as *competitive* situations. They will keep constantly in mind the differences between the purposes

and the symbols of education, and they will proclaim repeatedly their objection to any procedures which emphasize the symbols and obscure the purposes.

Research has made it clear that classroom teachers who set out to do these things find their work challenging and interesting. Children, under the direction of teachers who use the newer practices that result from recognition of individual differences in growth, concern for the individual, and analysis of the real purposes of education, do not lose in academic proficiency and do make substantial gains in personal-social development.

We have been discussing ways in which to improve the marking of pupils in terms of accomplishments in learning the subject-matter of the curriculum. Before considering personal-social development, we should pay some attention to devices that may be used to get information on which marks are to be based.

Standardized Tests

Investigations of the merits of many tests have indicated that there is great need of improvement in standardized tests. One authority recently stated the case when he said: "The millions of man hours devoted to preparing, administering, scoring, interpreting and defending mental measurements have yielded only a tiny dust heap of knowledge about language, thought, growth and meaning." Experts are aware of the many shortcomings in the available tests. There is no value in debating the issues here.

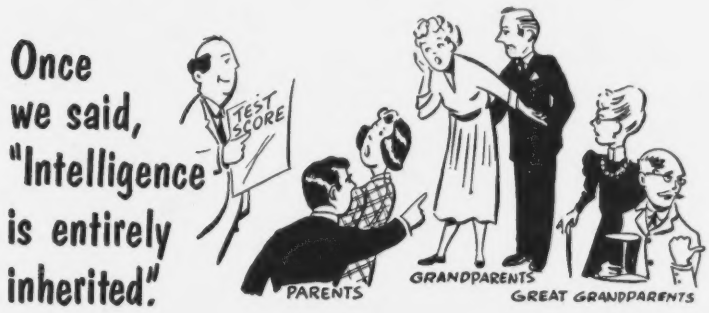
For most classroom teachers and school administrators the important thing is to look carefully before embarking on any elaborate standardized testing program. Evaluation of pupil progress for instructional purposes may be better if use is made of locally constructed devices that have been sharpened up by procedures invented by test makers and effective teachers. The great need is for tests flexible enough to deal with our own kinds of pupils, our own objectives, and our own school facilities.

Studies are being made of the influence upon test results of such factors as school size, teachers' salaries, teachers' training, and the availability of such facilities as libraries and science laboratories. Until the *evidence* from such studies is available, we must question the practice of comparing our pupils' perform-

ances with national norms except for general survey purposes by administrative officers. (The norms may have been developed on populations differing from our own in schools where instructional goals differ from ours.) Only rarely will a classroom teacher find a standardized achievement test score more helpful in the evaluation of his pupils' progress than the observational and diagnostic procedures that he can devise. At times the standardized test will fit the current learning situation and objectives, but for the day-to-day activities of the teacher they will seldom be pertinent and often are not informing with respect to instructional needs.

Use of Norms and Standards

If standardized tests are used, it is essential that we be clear about the term "norms." A norm tells us only the *typical performance* of the group on which the test was standardized. The fact that a particular group of pupils has scored above or below that norm does not reveal whether or not that class is doing superior or inferior work. To decide the quality of work, we would need to compare the purposes, facilities, curriculum, and instruction of our group with the norm group.



The business of comparing a child's performance with some vague "potential" based on scores of general intelligence tests is virtually outmoded. To say that a child is or is not working up to his ability when that ability is only a score on an intelligence test suggests erroneously that the test score is a highly

stable measure of some over-all potential that can be mustered for use in any subject at any time. We have in the past said that intelligence was "inherited" and would not change. We know now that this claim is a slim basis for blaming grandparents.

Let's look at it this way. Suppose that on a particular day a pupil makes a poor score. If we judged him as to whether or not he was "working up to his ability" by comparing future scores with one bad performance, we would be doing what we do when we contrast a golfer's worst performance with all his earlier scores and decide that he is not golfing up to his ability. The teacher who plays golf wouldn't like to have all his later achievements compared with one particular performance and the children don't like it either. And what about that impossible phenomenon, the child who is "working beyond his ability"?

In our discussion here we shall omit the ambiguous words *ability* and *aptitude*. Until research clears up the confusion as to whether ability and aptitude are things which are changing or things which do not change and until we decide whether they are goals to be achieved or determiners of goals, we should talk about them with great discretion.

Why should a child always be expected to work up to his ability even if the concept is valid? What child could stand up very long under such a program? Studies of evaluation procedures, in which a child's performance is always expected to be up to some vague potential, indicate that an objective of some schools is to keep pupils just short of the collapsing point.

Purpose of Essay Tests

For some years the essay test fell into disrepute when measurement enthusiasts pointed to the lack of validity and reliability in such tests. Essay tests have been so scorned that some teachers either have stopped using them or have felt guilty if they did use them. By so doing, they lost sight of the value of a carefully prepared essay test in showing how pupils can synthesize and analyze information. This value is admitted by measurement experts who resort to the use of essay examinations for candidates for advanced degrees. It is still true, of course, that badly and

hurriedly prepared essay tests have little merit and can give false impressions of pupil achievement.

Most objectively scored examinations require only simple recognition or recall. When the major purpose of an objectively scored test is to find out how much pupils can recognize and recall, as at times it may well be, the classroom teacher should use it. If, however, the teacher is concerned with pupils' achievements in recognizing relationships, expressing ideas, or analyzing and synthesizing information, he should not hesitate to use well-prepared essay examinations. Scoring them, it has been found, will be significantly improved if the questions have been worded so that the teacher can prepare in advance a list of important points to be expected in the pupils' answers. Also, scoring results are likely to be better if the teacher reads the answers to the first question on all the papers, then the answers to the second question, and so on rather than reading all the answers of each student consecutively.

Teacher-Made Objective Tests

The so-called "objective test" is really a subjectively constructed test that is objectively scored. The actual writing of test items is a subjective process. The author of an objectively scored test must *choose* the form of items to use, must *decide* on the materials he will sample, must *make judgments* about whether or not an item is worthy of inclusion, and must *select* among scoring schemes. These are among the subjective procedures that call for study by those planning to use an objectively scored test; no objective scoring system can ever make up for faulty subjective decisions made during the construction of the test.

Teacher-made, objectively scored tests may involve simple recall of information by requiring the pupil to complete sentences or to fill in blanks. Teachers may also use recognition tests of the true-false, multiple choice, or matching varieties. At first glance, these tests seem easy to prepare but the construction of a really good objectively scored test is a difficult undertaking if certain pitfalls are to be avoided and the full merits of these tests are to be obtained.

True-false tests are popular because they *seem* to be easy to construct, because a great many items can be covered quickly, and because they are easy to score. This type of test encourages bad practices when the items are just lifted from textbooks. In this form they encourage pupils to memorize isolated bits of information, to do wild guessing, or even to get false ideas from incomplete items. Students who *understand* rather than memorize are frequently penalized by true-false items; often the more a pupil knows about certain topics the more difficult it is to answer the item as either true or false.

If the true-false type of test is to be used, it should be limited to such items as: "There are 12 inches in a foot." Here there can be no doubt about the item being true or false and the teacher can be sure that knowing the item in isolation is important. There are not many items that are of value in isolation; most of them can be appraised thru better ways of testing. The true-false type of test should be used sparingly.

The best type of objectively scored test is composed of multiple-choice items in which a statement or question is followed by several possible answers. Its excellence lies in the discrimination required of the pupil; he must use the information learned to select the best answer. The items in a multiple-choice must be prepared with care since it is difficult to get enough plausible choices and hard to avoid giving clues to the correct answers. A good form of item is one giving certain kinds of information and asking the pupil to state his conclusions from the data and to choose from a number of reasons those which support his conclusions.

What has been written here about the value of the true-false and multiple-choice types of tests should be read again with the earlier discussion of the essay test. Research shows that reliability (really consistency in responding) of the objectively scored tests is generally higher than the reliability of the essay test and that more material can be covered in a shorter time. There is real danger, however, that if we limit ourselves to these newer devices, we teachers may lose sight of some of the most important goals in education. We may tend to encourage the development of "multiple-choice and true-false minds" that are unable to synthesize information and to understand the meaning of facts.

Educational testing should do more than measure the temporary visual suspension of what is in our textbooks.

ASSESSING PERSONAL-SOCIAL DEVELOPMENT

We have been considering, for the most part, the evaluation of the so-called "intellectual development" of the pupil. We turn now to the problem of assessing development in personal-social characteristics. Our separation of topics is for convenience of discussion and distribution of emphasis rather than because the topics are really separate or that one is more important than the other. Research studies have shown that, while the child is learning the Three R's, he is also learning the Three C's (character, cooperation, and courtesy) and many other characteristics that go to make up what is considered personal-social development.

Many Questions Unanswered

In the discussion of the assessment of school achievement we suggested that research had not given us the answers to all questions and it will be evident, in the evaluation of personal-social development, that there are many more unanswered questions. This area contains many factors often called "the intangibles," interpreted by some as meaning that teachers cannot evaluate progress toward school objectives in this area. Currently that interpretation comes close to the fact, but research has developed some techniques that can be used fairly effectively if great care is taken in their selection, construction, and interpretation.

Checklists and Rating Scales

A commonly used procedure for appraising "intangible" learnings is the checklist. Often classroom teachers indicate a pupil's development by checking or putting a symbol after a word or short description of some aspect of personality. Thus, the teacher may be asked to check whether a child's "citizenship" is excellent, good, fair, or poor. He may be asked to indicate which of a number of briefly defined characteristics placed in supposed order

of excellence best describes the behavior of the pupil as he has observed it. Rating and checking of this kind is likely to be a compilation of errors, for the procedure is a relic of the dark ages of psychology. The procedure would be discarded if it did not *seem* to accomplish its purpose in a very short time.

Actually such ratings may be definitely harmful since research has shown that among their other limitations they (a) suggest that certain characteristics are equally desirable in equal amounts for all individuals at all times; (b) encourage generalization about a pupil's characteristics beyond what was actually observed by the rater; (c) encourage the making of comparisons of pupils who are quite different and who have had unequal environmental opportunities; (d) assume that teachers can observe behavior, such as cooperation, sort it into units on a scale, and allot values to it; (e) suffer from "halo" effect, that is, the teacher who rates a pupil high in one characteristic tends to rate him high in others or vice versa; and (f) usually suffer from inadequate definition of the terms to be rated so that what is satisfactory to one person may be very unsatisfactory to another.

The foregoing are the major limitations but there are many minor ones—so many in fact that it is difficult to see why checklists and rating scales continue to play a major part in evaluating and reporting about pupils. Perhaps their use continues because they seem to be effective in doing an objectionable task in an area where teachers feel insecure.

The basic difficulty lies in the "judgment" aspects of rating. In accepting a rating scale, one also accepts the philosophy of the person who constructed it. Consider, for example, the following rating scale in which the statements are supposed to be arranged in descending order of excellence under the heading "responsibility":

1. Finds additional unassigned tasks after completing assignments
2. Helps others in the class after completing own assignments
3. Completes only the work assigned to him
4. Needs occasional prodding to get his assignments done
5. Needs unusual amount of prodding to do his assignments.

Although in this list more than the usual effort has been made at definition, there still are problems of understanding and philos-

ophy. If you use the scale in this form, you put doing additional tasks for yourself ahead of helping others. You put a premium on being the kind of "eager beaver" who may be proceeding rapidly toward a nervous breakdown or a case of ulcers. It is possible that item 3 in the list is the best of these behaviors; that is, you should do what you have to do (as well as you can) and then go fishing. Research on individual differences suggests that it is unrealistic and unwise to try to place pupils on the same scale without considering their unique circumstances and situations. Perhaps as our knowledge of individual differences increases, the common rating scale or checklist will become obsolete.

Observations vs. Measurement

As any science develops, the processes of observation always precede those of measurement. In our study of the personal-social behavior of growing human beings we are still largely in the observation stage. Techniques for observing pupils range from the drawing of sociograms of a class to the collecting of anecdotal records and simple reporting of behavior descriptions.

As we consider descriptive procedures, we should try to distinguish between those devised by research workers to sharpen our observations periodically and those which can be used as regular procedures. Thus, a sociogram demonstrating the social relationships within a class may help us to remember that social relationships are important and may help the classroom teacher to recognize accepted, rejected, and neglected children. However, the drawing of as many sociograms as shifting class relationships require soon becomes a heavy task. Some of us may reject the technique because it simply elaborates the obvious to the classroom teacher who is alert in observing social relationships within his classes.

Use of Anecdotal Records

The anecdotal record method of observing, interpreting, and reporting pupil progress has received a great deal of attention during the past two decades. Like other techniques that are newly

introduced it became a fad, reached its zenith in popularity, and has now found its place among other techniques. It has some value when used properly, timely, and with caution. The "anecdotes" in a record are descriptive accounts of episodes or events in the daily life of the pupil, with some interpretation of their significance in his development. All classroom teachers observe these events and, unless an attempt is made to record them, they must depend upon memory to evaluate a pupil's growth. It is true, also, that if incidents are recorded and interpreted, the one making the record is more likely to be diligent in observation and more serious in his efforts to understand a pupil's behavior.

In practice, anecdotal records have been concerned more with social relationships than with subjectmatter accomplishments; they can be of value in both areas. Thus, observation of a pupil in the classroom may reveal vigor or lassitude of response, variation from usual behavior under specific stimulation, a tendency to go beyond minimum requirements, attempts to improvise, reactions to authority, and relative degrees of zeal or apathy in response to various activities. The following anecdotal records and a summary (obtained and reported by a specialist in child development) indicate how they may be used to highlight social relationships.

October 1—Asked if he could go around the rooms to collect the milk bottles. He had this job for most of last year.

October 16—Went to cupboard, distributed crackers and milk without being told. As a rule he does not assume any responsibility. He usually waits until he is told or allows someone else to take the lead.

October 25—Talked to the class today about a radio program he had heard at a friend's house. This was one of the rare times he has spoken to the class; generally very quiet.

November 6—Finished a model airplane he has been working on. He has shown much interest here.

December 3—Read a comic magazine several times this morning. It is difficult for him to take an interest in class reading.

December 5—Did oral reading for the group during reading period. Did not read loud enough for the others to hear. He doesn't like to be in front of other people where he is the center of attention.

A brief summary and interpretation of these anecdotes is shown in the following statement: "These are the anecdotes on a pupil who finds it difficult to discover any school task he is capable of performing. The anecdotes as a whole give the impression of the pupil's willingness to participate where he is able, a normal interest in the adventures of comic magazines but not in school material, and a dislike of doing things in front of others. Each anecdote may be minor, yet the series gives a clearer picture of the pupil's responses to factors and situations in the classroom."

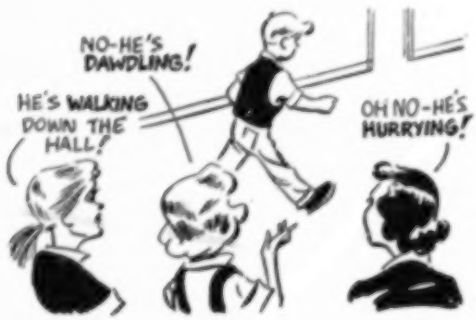
Limitations of Anecdotal Records

People who have appraised the written anecdotal record method are usually in favor of it, but the claims to objectivity in the method are greatly exaggerated. If a classroom teacher reports only that Jimmie "*walked* down the hall," it is said that the statement is objective. The *choice* of the word *walked,* however, involves a good deal of subjectivity. Another observer of Jimmie under the same circumstances might well have said he *dawdled* down the hall, still another that he *strode* and still another that he *hurried.* Research into the validity of the testimony of observers and the data obtained by specialists in word meaning raises considerable doubt as to the objectivity of observational techniques, even when they are used by trained observers.

Enthusiasts for the anecdotal record recommend that many records be made so that patterns of characteristic behavior may evolve. In our enthusiasm for collecting many anecdotes we should not forget that the affections of school personnel can be alienated away from a good evaluating and reporting program by insistence upon a large number of anecdotes on a specific number of pupils in prescribed form at a given time. Rather, the classroom teacher should be asked to report behavior which is consistent with (or significantly different from) the activities of the

pupil as he has observed and understood him. And very important, a workable plan includes personnel to do the clerical work and the summarizing of the anecdotes. Something reasonable, as outlined, would make anecdotal records a valuable supplement to records of a pupil's development.

In anecdotal records—words must be used with care



It has been amply demonstrated that a technique, called *The Method of Behavior Description*, can provide valuable evidence of pupils' development in the areas usually covered by rating scales, checklists, and anecdotal records. The entries in this "behavior description method" resemble those in checklists and rating scales, but there is a basic difference—the user's belief in the importance of individuality and the importance of individual differences. First, those who plan to use the method must prepare carefully defined descriptions of the pupil's behavior, (sometimes similar to those used in rating scales). Classroom teachers, and others with sufficient opportunity to observe the pupil, place symbols indicating their relationship to him beside the description which best fits him. The chart on page 21 illustrates the method.

This description indicates that Mary felt secure and was well accepted in her English and music groups in Grade VII, but had begun to show some anxiety about her relationship to her peers in those classes in Grades VIII and IX. Something happened to cause the girls in the home economics class to treat her with indifference in two upper grades. The classroom teacher, wanting to help Mary, would seek to uncover the events leading to the

damage. In addition to the abbreviations placed opposite the descriptive items some teachers will want to add supplementary notes or explain what lies behind the appraisals they have given.

Pupil: Mary Anderson—Junior High School

Descriptions	Grades		
	VII	VIII	IX
Appears to feel secure in and is accepted by groups of peers	E, MU		
Appears to feel anxious about her standing in her groups		MU, E	MU, E
Wants to belong to groups but is generally treated with indifference		HE	HE
Withdraws from peers so much that she is not fully accepted			
Characteristics of her person or behavior cause rejection by her group			

(E = English; MU = Music; HE = Home Economics)

Description vs. Rating

The difference between this descriptive procedure and rating is just that the describers try *to summarize* what has been observed while raters attempt *to judge* the quality of the observed behavior. There is no implication in the "behavior description method" that any particular kind of behavior is best for any one child at a particular time. The technique admits the well-known fact that a child's behavior may vary in different situations and under changing influences. Thus, though each reporter makes a correct description of what he observes, the reports about an individual may differ greatly at any given time. The plan allows for the possibility that differences in the descriptions of various observers may be as significant as the differences they report.

It must be emphasized that there is no implication of goodness or badness in the use of the term *behavior*.

Instead of requiring a perfunctory rating of personality twice a year, a practice that classroom teachers dislike and if possible avoid, the behavior description method proposes that teachers be encouraged to make continuous observations of their pupils with respect to the defined characteristics and to record their descriptions at such times as are decided upon. Duplicated sheets of the definitions of characteristics are furnished to the teachers so that they can make their descriptions of the pupils with the definitions before them, and without being influenced by each other's observations. The descriptions can be entered on sheets of class lists with the characteristics used as headings across the top of the page. Abbreviations and numbers for types make such a form simple to prepare. The descriptions are transferred from the class list to the central record card, thus making a picture of the pupil as seen by all his teachers. When significant notes accompany a description, they can be entered on the record card beside the definitions. If teachers study the form and the definitions of behavior at the beginning of the school year and agree to make the descriptions upon the basis of carefully considered evidence, the descriptions are likely to be valid.

Personality and Interest Questionnaires

The word *questionnaire* in the heading of this section has been used advisedly because there is no such thing as a personality or interest *test* in any legitimate use of the word *test*. The terms *personality test* or *interest test* may have been used by some publishers to appeal to the psychometrically innocent school personnel. The term has been accepted by them probably because they use the format and the paraphernalia (norms, validity reliability quotients and standardized directions) of tests of achievement. Despite all the propaganda and paraphernalia the classroom teacher and all others who use them must be warned that they are simply questionnaires and subject to all the limitations of the questionnaire technique. After thorough examination of the research literature and the so-called personality and interest tests themselves this author recommends that it is desirable to have a

period of years during which test authors and publishers can establish satisfactory evidence of the validity, forecasting efficiency, and diagnostic value of personality questionnaires. Most of these forms elaborate the obvious; others make more cloudy what is already unclear. Certainly the mass administration of these devices is an unwise use of time and money. If, despite all their limitations, they are to be used at all, they should be employed only for that rare case where every possible device is employed in the hope of throwing some light on a very difficult situation.

Research has demonstrated that pupils can fake their answers in personality and interest inventories. The answers called for require only superficial snap judgments about which the subject may be ill-informed or uninformed; the pupil being "tested" is often required to make choices where he has no real choice or concern; there is inadequate opportunity for the pupil to indicate the depth, stability, or permanence of his attitudes and enthusiasms. The personality questionnaires rely on self-judgments, probably the least valid of all kinds of appraisals. Their facile profiling of scores implies that something valid has been measured so reliably that differential diagnosis is possible. At most, the personality inventory lists symptoms which can be identified equally well and more economically by observations. They assume an atomistic structure of personality and interest patterns and they lump these atoms and name them by fiat with little relationship to the true nature of personality.

It should be kept in mind also that the study of the personality of a child is not a simple process accomplished by answering a few questions in a 20- to 30-minute period. Classroom teachers, who spend much time with children and youth, know that the understanding of personality problems is not likely to come from the application of instruments of this type. We still lack incontrovertible proof that the instruments provide adequate facts.

Personal Documents and Projective Technics

Most questionnaires and tests may be criticized as so fully developed in structure and guide lines that the pupil has little opportunity to reveal the things that are important to him. Devices that avoid this difficulty (and sometimes go too far in the

opposite direction) are sometimes called "personal documents and projective techniques." Both these devices allow the pupil to express himself in the areas of his greatest concern and to give as much detail as he wishes. By making a full, unguided response, the child may reveal attitudes, desires, interests, and problems that will not be disclosed by more formal devices and methods.

Under the heading "personal documents" we may include autobiographies, diaries, or other written materials. Also, by stretching the definition somewhat, we may include drawings, constructions, performances, or oral recitations. Research shows that careful analysis of a pupil's productions may reveal clues about his reaction to school work and the level of his personal-social development. We may observe his clarity in expressing himself, make a systematic, complete evaluation of his writing, and assess his use of language forms. From such evidence we may get *indications* of interest; *clues* about social attitudes; *hints* of his tendency to try to make a good impression when given the chance; *clues* about his tendency to rationalize, compensate, or project; and *inferences* about persons or circumstances that have influenced him.

Since these are only hints, clues, indications, and inferences, the so-called "personal documents" approach cannot be used in isolation. The child forgets or omits; as observers, we may misinterpret. These limitations demand extreme caution in the use of personal documents. Their merit lies in the fact that they *may* reveal the "inside half" of the child's life. In our strenuous search for "objective" evidence, this half of personality may be overlooked and yet it must be examined. No classroom teacher can assume, because there are no *outward* manifestations of conflict, frustration, or disturbing feelings, that none exists.

Projective techniques also attempt to get at the inner half of personality. They are based on the premise that when an individual responds to an *ambiguous* stimulus, he is apt to expose his own personality as much as the phenomenon to which he is attending. Thus, a child may be asked to tell what he "thinks about" inkblots, vague drawings, or sounds. He may be given rather vague instructions to draw a person or to tell what is happening in a picture that is subject to many interpretations. Experience shows that in these explorations he may reveal latent

needs, images, and sentiments which he would be unwilling or unable to express in answer to direct questioning.

The "projective idea" may be used effectively by classroom teachers. They can use such stimuli as pictures, incomplete stories, and unfamiliar music. While doing so, they may observe the pupil's responses with respect to enthusiasm, apathy, choice of language, and the extent to which he identifies himself or others with the stimulating situation. Used in this way, projective methods provide some evidence about a pupil which, when combined with other information, may provide suggestions for next steps in his treatment.

Not in Isolation

Throughout this long discussion of "How Should We Evaluate?" the reader must have observed much reiteration of statements to the effect that no one of the many techniques considered has sufficient merit to recommend its use in isolation. As research on the nature and needs of growing human beings accumulates, it becomes clearer that evaluation is a complex process requiring the use of many methods.

Perhaps the most serious error that classroom teachers could make in the area of evaluating a pupil's progress would be that of becoming converts to one or even a few theories about behavior or methods of evaluating and reporting about it and of using them exclusively. The evaluative process becomes a process of sifting evidence, analyzing situations, bringing together isolated facts, and thinking over again and again the apparent significance of the facts. The task is never finished as long as the pupil is in the school, for both he and the school are changing and each year new methods of evaluation are developed.

WHEN SHOULD WE EVALUATE?

We may emphasize again that evaluation is a continuous process and an integral part of instruction. The final examination at the end of a course on which the student passes or fails the year's work is fast disappearing from public-school practice. Many tests that are given at the end of the school term are now

used in planning the work for the coming year. They are not used to fail or pass pupils but to determine their readiness for next steps and to diagnose difficulties so that remedial treatment can be provided.

Evaluation, A Year-Round Procedure

But these purposes are not solely end-of-the-year purposes. Evidence of accomplishment, growth, and development is needed throughout the whole school year so that classroom teachers may see what progress is being made toward the objectives. Readiness for next steps does not always emerge at precise days or hours indicated by the calendar or the clock. We must be alert to recognize readiness when it appears lest we lose the golden opportunities when pupils are eager to learn. As classroom teachers, we must recognize lack of readiness, too, lest we require the pupil to study something too soon and actually set him back so that he does not learn it at the usual time. With a plan for continuous evaluation we are more likely to identify readiness (or lack of it) than we are when we depend upon chance.

Evaluation as Diagnosis

Diagnoses of difficulties which may arise at any time cannot await annual examinations. Progress must be observed and recorded when it occurs or faulty memory may obscure it. New patterns of behavior or retrogressions to past patterns need to be evaluated so that action may be taken upon them at the right time. Evaluation cannot mean just the administration of final tests for promotion purposes alone. Evaluation should provide a moving picture rather than a snapshot of pupil development.

WHO SHOULD EVALUATE?

Ultimately the responsibility for evaluation falls upon those who teach. Parents may assist and pupils may contribute to the process. In situations where several classroom teachers work with

the same pupils, evaluation may be the cooperative product of several teachers, pupils, parents, counselors, homeroom instructors, visiting teachers, physicians, school nurses, employers, or, in fact, any persons who have had sufficient opportunity to observe the pupil. Many studies tell us that the behavior of a pupil may vary as circumstances change so that what one person observes, another does not see. Both may have seen different facets of the pupil's behavior, apparently conflicting, which are truly descriptive of that variability in his behavior which may be most important in his guidance.

Pooled data have generally (but not without exception) been found to be more valid than information from single sources. When many persons are informed about objectives and the need for securing evidence of progress toward them, and when they learn about methods of securing such evidence, they are likely to sharpen their observations. When they have opportunities to pool their facts and discuss them, they are more likely to see the whole child than when they have no responsibility for evaluation or can delegate it to another person. All these findings suggest, then, that evaluating and reporting a pupil's development will be more effective if all those who have had sufficient opportunity to observe him have the responsibility of reporting their observations. Procedures for reporting will be discussed in the following sections.

REPORTING PUPIL PROGRESS

Collecting, coordinating, and reporting information about a pupil's progress for parents, pupils, and the next teacher is a complicated task that requires much more attention than has commonly been given to it.

Typical Report Cards

An analysis of a small sample of report cards revealed that in the area of personal characteristics alone (excluding terms used to describe accomplishments) some 260 trait names were employed. Most of the report cards used such terms as *cooperation, industry, citizenship,* without any definition and in most cases

the record consisted of a single check mark after a trait name and under such headings as superior, average, fair, and poor. The remaining parts of the reports contained letter grades or marks and a statement about attendance and tardiness.

These characteristics appear to be typical of the report cards sent home about four times a year to share information fundamental to a cooperative relationship between the classroom and the home. It is unlikely that they accomplish that objective. Recognition of the limitations of such reports results in numerous revisions and many attempts to provide other methods for providing the needed information. Since many report cards also require the use of cumulative records within the school, it may be well to discuss the nature of such records before we consider the improvement of reporting.

Cumulative Records

The cumulative record card (or folder) has achieved much prominence within the past quarter of a century. Essentially it is a record of pupils' academic achievements, test performances, interests and activities, health and physical development, usual behavior, and interview reports compiled at stated periods and accumulated over a span of years. Occasionally it contains records of treatment given, results of home visits, special actions taken, or recommendations made. Also listed may be census data such as places of residence, age, occupations of parents, numbers of brothers and sisters, and brief home descriptions. Commonly these data are compiled on a cardboard form or in manila folder which is used for reference when a pupil is to be discussed or reports are to be made. The chief merit claimed for the cumulative record is that it presents a rather complete picture of a pupil brought together from many sources over a long enough period to portray his development. When properly constructed and efficiently kept, a cumulative record becomes an effective procedure for recording pupil development. It is an indispensable element in an efficient school.

The cumulative record should not be thought of primarily as a bookkeeping device. Although much of its content and form is statistical, its function goes far beyond a reference place for fac-

tual material. The statistical data, combined with data from other sources to make the record, have been shown to be helpful for the following purposes:

1. They form a basis for current guidance.
2. They preserve guidance material so that it can be used with continuity over a term of years in as many schools as the pupils attend.
3. They help the school staff to give advice about a pupil's further education or entrance into an occupation, and to furnish the material for reports to other schools or agencies on his readiness for the new experience and his probability of success in it.
4. They provide information for use in cooperation with parents, either through written reports or by means of interviews.

Teachers' Attitudes Toward Reporting Pupil Progress

Many classroom teachers dislike writing reports and making records. The task is often regarded as extra clerical work that is to be done at certain times, such as at the end of the school year, when there are many other stresses. Many errors occur, feelings may be hurt, and the entire process is regarded as, at best, a necessary evil. To some teachers, however, the making of records and reports is not a dreary, routine, clerical task but a challenging study of growing boys and girls. The difference between these attitudes toward recording and reporting is that the first fails to look beyond the immediate task at hand, sees only part of the job and even sees that incompletely. The second tries to assist in putting together the many-sided puzzles of individual achievements, interests, personalities, attitudes, ambitions, health, and home circumstances and seeks to assist in the solution of these puzzles so that happiness and success for the pupil are more likely than failure and disillusionment.

Teacher-Parent Conferences

It is not enough to *have* cumulative records. They must be *used* in reporting to parents, future teachers, and employers. The records must not be dead; they must come alive under the interpretation that competent teachers can give them.

We recognize that parents are familiar with the old-fashioned report and they like them because the familiar is comfortable. Parents tend to distrust anything that *seems* less definite than the letter grade or numerical mark and they feel more secure with something that seems to be more objective. If they are to use and appreciate reports that come from systematic cumulative records and elaborate analysis of child behavior, a well-planned program for informing them must be attempted. In some schools this has been done by appointing parents to school committees on cumulative records and reports and in others by public information programs in the parent-teacher association, in the press, and on the radio.

In many communities parents of elementary-school children are invited to come to the school several times a year to discuss their children's progress in conference with the teachers. The discussion revolves around the past development portrayed in the cumulative record and plans are laid for cooperation in working toward mutually acceptable goals. Although technical research does not indicate whether or not this is a desirable practice, it seems to be a common-sense method of sharing responsibilities.

New Reporting Procedures Needed

The processes of reporting pupil progress are in a state of flux. There is demand for information that will tell parents and others with definiteness where their children are showing strengths or weaknesses as judged by normal expectations of children of their ages and opportunities. There is also demand for information that describes a pupil's progress in a way analytical enough to give helpful guidance and to indicate the pupil's likelihood of success in continuing to work in certain fields, both in later years in school and in advanced institutions. There is still need for the invention of a way to direct the minds of pupils, parents, and classroom teachers away from marks toward the fundamental objectives of education. Such reports need to show appreciation for the poorest pupils' good qualities while the best pupils' weaknesses are pointed out. When reports can do these things and also add recommendations of ways in which pupils can be helped to overcome weaknesses and use strengths more effectively, they can become potent tools in the improvement of schools.

A FINAL WORD

Some will be disappointed at the numerous limitations that research has revealed about commonly used techniques for evaluating and reporting pupil progress. Disappointment may lead to discouragement, feelings of frustration, a general sense of futility, and a decision to do nothing about an apparently hopeless situation. There is, however, another way to look at a disappointing situation and it is this view that professional teachers will take. They will see it as a challenge to their ingenuity, skill, and good judgment. They will realize that research has indicated many pitfalls but they will note that it has illuminated numerous pathways to improvement. They will recognize that research has provided alternatives to uncritical passive acceptance of programs of evaluating and reporting that accomplish less than they might because they are static. Those alternatives are for classroom teachers and research workers to appraise continuously our evaluation programs and to persist in their attempts to determine the effectiveness of methods for reporting pupil progress. As these things continue, we will see that there is much merit in many of the time-proven instruments when sharpened by the techniques described in this bulletin. It will be observed also that the vigorous attempts of research investigators to develop more highly effective tools show enough promise to warrant an optimistic view of the future.

SELECTED RESEARCH REFERENCES

Books

1. Association for Supervision and Curriculum Development. *Evaluation as Feedback and Guide*. 1967 Yearbook. Washington, D.C.: the Association, a department of the National Education Association, 1967. 282 pp.
2. Beauchamp, George A. *Basic Dimensions of Elementary Method*. Second edition. Boston: Allyn & Bacon, 1965. pp. 298-316.
3. Blount, Nathan S., and Klausmeier, Herbert J. *Teaching in the Secondary School*. Third edition. New York: Harper & Row, 1968. "Reporting Pupil Progress," pp. 417-42.
4. Bossing, Nelson S., and Crammer, Roscoe V. *Junior High School*. Boston: Houghton Mifflin Co., 1965. Chapter 12, "Evaluating and Reporting Student Progress," pp. 283-314.

SELECTED RESEARCH REFERENCES

5. Clark, Leonard H., and Starr, Irving S. *Secondary School Teaching Methods.* Second edition. New York: Macmillan Co., 1967. Chapter 16, "Marking and Reporting to Parents," pp. 380-406.
6. Cooper, John E. *Elementary School Principalship.* Columbus: Charles E. Merrill Books, 1967. pp. 148-56.
7. Kendred, Leslie W., and others. *Intermediate School.* Englewood Cliffs, N.J.: Prentice-Hall, 1968. "Reporting Pupil Progress," pp. 275-81.
8. Mercer, Jane R. *Labeling the Mentally Retarded.* Berkeley: University of California Press, 1972. (In press.)
9. National Education Association, Research Division. *Marking and Reporting Pupil Progress.* Research Summary 1970-S1. Washington, D.C.: the Association, 1970. 53 pp.
10. Strang, Ruth. *How To Report Pupil Progress.* Chicago: Science Research Associates, 1955. 47 pp.
11. Tweksbury, John L. *Nongrading in the Elementary School.* Columbus: Charles E. Merrill Books, 1967. Chapter 6, "Reporting Pupil Progress," pp. 100-22.
12. Van Til, Williams; Vars, Gordon F.; and Lounsbury, John N. *Modern Education for the Junior High School Years.* Second edition. Indianapolis: Bobbs-Merrill Co., 1967. "Reporting," pp. 532-40.
13. Wrinkle, William L. *Improving Marking and Reporting Practices in Elementary and Secondary Schools.* New York: Holt, Rinehart and Winston, 1947.

Periodicals

1. Austin, Mary C. "Report Cards and Parents." *Reading Teacher* 18: 600-63: May 1965.
2. Center for the Study of Evaluation. "Characteristics of Standardized Tests as Evaluation Instruments." *UCLA Evaluation Comment,* September 1971.
3. Christensen, Donald J. "Effect of Discontinued Grade Reporting on Pupil Learning." *Arithmetic Teacher* 15: 724-26; December 1968.
4. Chansky, Norman M. "How Invalid Are Marks Assigned by Teachers?" *Journal of Educational Psychology* 43: 218-19; April 1952.
5. Climo, James. "Pass/Fail at Longmeadow." *Clearing House* 43: 341-43; February 1969.
6. Cureton, Louise Witmer. "The History of Grading Practices." *Measurement in Education,* Vol. 3, May 1971.
7. The National Elementary Principal. "Reporting to Parents." *The National Elementary Principal* 65: 6-51; May 1966.
8. Fensch, Edwin A. "Tone Down the Bell Curve." *Clearing House* 40: 469-71; April 1966.
9. Flinker, Irving. "How Do We Rate?" *Clearing House* 42: 245-47; December 1967.

10. Garth, Warner D. "Marks–How Much Do They Mean?" *PTA Magazine* 63: 2-4; April 1969.
11. Gillcrist, William A., Jr. "Study of Grading in Arithmetic." *Education* 86: 177-81; November 1965.
12. Haskell, John M. "Pass–Fail? A System Worth Trying." *Clearing House* 42: 72-73; November 1967.
13. Halliwell, Joseph W. "Dangers Inherent in Correlating Averages." *Journal of Educational Research* 55: 327-29; April 1962.
14. Klemm, Eugene. "Parents and Report Cards." *PTA Magazine* 61: 24-26; October 1966.
15. Lange, Phil C. "Taking the Stress Off Grades." *PTA Magazine* 62: 19-21; October 1967.
16. LeFevre, Carol. "Face to Face in the Parent-Teacher Conference." *Elementary School Journal* 68: 1-8; October 1967.
17. Link, Frances R. "To Grade or Not To Grade." *PTA Magazine* 62: 10-12; November 1967.
18. McGuire, Brian P. "Grading Game." *Today's Education* 58: 32-34; March 1969.
19. Mann, Lester, and others. "Comparison of Formal and Informal Reporting Systems in a First Grade Population." *Journal of Educational Research* 60: 75-79; October 1966.
20. Marshall, Jon C. "Evaluation: Does It Mean Anything?" *Clearing House* 42: 535-38; May 1968.
21. Mikaelian, Sam, and Thompson, Ronald B. "Evanston Tried Computer Report Cards." *Instructor* 77: 129; March 1968.
22. NEA Research Bulletin. "Report to Parents." *NEA Research Bulletin* 45: 51-53; May 1967.
23. NEA Research Memo. "School Marks." *NEA Research Memo.* August 1964.
24. Riban, David M. "Grade-Weighting System Investigated." *Bulletin of the National Association of Secondary School Principals* 50: 94-98; November 1966.
25. Rosenthal, G. H. "Where Did the System Fail?" *Clearing House* 41: 558-61; May 1947.
26. School Management. "If You're Thinking About Dropping Report Cards." *School Management* 11: 104-109; April 1967.
27. Solley, William H., and others. "Grading in Physical Education." *Journal of Health, Physical Education, and Recreation* 38: 34-39; May 1967.
28. Terwilliger, James S. "Self-Reported Marking Practices and Policies in Public Secondary Schools." *Bulletin of the National Association of Secondary School Principals* 50: 5-37; March 1966.
29. Trow, William C. "On Marks, Norms, and Proficiency Scores." *Phi Delta Kappan* 48: 171-73; December 1966. Condensed: *Education Digest* 32: 35-37; March 1967.
30. Yauch, Wilbur A. "School Marks and Their Reporting." *NEA Journal* 50: 50, 58; May 1961.
31. Warren, Jonathan R. "Current Grading Practices." Research Report Number 3. *American Association for Higher Education.* January 15, 1971.
32. White, Mary A., and Boehm, Ann E. "Child's World of Marks." *NEA Journal* 57: 12-13; January 1968.

FILMED FROM BEST AVAILABLE COPY

WHAT
RESEARCH
SAYS SERIES